

Fisher information as a performance metric for locally optimum processing

Fabing DUAN, François Chapeau-Blondeau and Derek ABBOTT
 Institute of Complexity Science, Department of Automation Engineering,
 Qingdao University, Qingdao 266071, PR China
 Laboratoire d'Ingénierie des Systèmes Automatisés (LISA),
 Université d'Angers, 49000 Angers, France.
 Centre for Biomedical Engineering (CBME) and School of Electrical
 & Electronic Engineering, The University of Adelaide,
 Adelaide, SA 5005, Australia

November 28, 2011

Abstract

For a known weak signal in additive white noise, the asymptotic performance of a locally optimum processor (LOP) is shown to be given by the Fisher information (FI) of a standardized even probability density function (PDF) of noise in three cases: (i) the maximum signal-to-noise ratio (SNR) gain for a periodic signal; (ii) the optimal asymptotic relative efficiency (ARE) for signal detection; (iii) the best cross-correlation gain (CG) for signal transmission. The minimal FI is unity, corresponding to a Gaussian PDF, whereas the FI is certainly larger than unity for any non-Gaussian PDFs. In the sense of a realizable LOP, it is found that the dichotomous noise PDF possesses an infinite FI for known weak signals perfectly processed by the corresponding LOP. The significance of FI lies in that it provides an upper bound for the performance of locally optimum processing.

1 Introduction

Fisher information (FI) is a measure of the minimum error in estimating an unknown parameter of a distribution, and its importance is related to the Cramér-Rao inequality for unbiased estimators [1, 2]. By introducing a location parameter, the de Bruijn's identity indicates that the fundamental quantity of FI is affiliated with the differential entropy of the minimum descriptive complexity of a random variable [1]. Furthermore, in known weak signal detection, a locally optimum detector (LOD), as an alternative to the Neyman-Pearson detector, has favorable properties for small signal-to-noise ratios (SNRs) [3]. With sufficiently large observed data and using the central limit theorem, it is demonstrated that the LOD is asymptotically optimum and its asymptotic efficiency is upper bounded by the FI of the distribution [2]–[5].

However, the fundamental nature of FI is not adequately recognized for processing known weak signals. To extend the heuristic studies of [1]–[6], in this paper, we will theoretically demonstrate that, for a known weak signal buried in additive white noise, the performance of a locally optimum processor (LOP) is completely determined by the FI of a standardized even probability density function (PDF) of noise. We show this for three signal processing case studies: (i) the maximum SNR gain for a periodic signal; (ii) the asymptotic relative efficiency (ARE) of a LOD for signal detection; (iii) the best cross-correlation gain (CG) for an aperiodic (random) signal transmission. Moreover, for estimating an unknown parameter of a weak signal, the minimum mean square error

of the unbiased estimator can be reduced to a straightforward form expressed by the FI of the distribution. The physical significance of FI, resulting from the reciprocal of FI delimiting the minimum mean square error of unbiased estimators, provides a upper bound of the performance for locally optimum processing. It is well known that a standardized Gaussian PDF has minimum FI of unity [2]. As a consequence, for any non-Gaussian noise, it is always possible to achieve the performance (SNR gain, ARE or CG) of a LOP larger than unity for the three considered situations. In the sense of a realizable LOP, an example of a Gaussian mixture noise PDF is investigated [9]. It is found that arbitrarily large FI can be achieved by the corresponding LOP, and even when the noise is dichotomous noise associated with infinite FI.

2 Fisher information measures performances of locally optimum processing

Since the known weak signal might be periodic or aperiodic, three signal processing cases are illustrated for exploring the significance of FI in locally optimum processing.

2.1 SNR gain for periodic signal processing

First, consider a static processor with its output

$$y(t) = g[x(t)], \quad (1)$$

where the nonlinearity g is odd [6] and the input is a signal-plus-noise mixture $x(t) = s(t) + z(t)$. The component $s(t)$ is a known weak periodic signal with a maximal amplitude A ($0 \leq |s(t)| \leq A$) and period T . A zero-mean white noise $z(t)$, independent of $s(t)$, has an even (symmetric) PDF f_z [4, 6] and the root-mean-square (RMS) amplitude σ_z (if it exists, or it is a scale parameter.). A family of even (symmetric) PDFs is frequently encountered in practical signal processing tasks [2, 4, 5, 6]. In the case of $A \rightarrow 0$, we have a Taylor expansion around z at a fixed time t as

$$y(t) = g[z + s(t)] \approx g(z) + s(t)g'(z), \quad (2)$$

where we assume the derivative $g'(z) = dg(z)/dz$ exists for almost all x (similarly hereinafter) [2, 5]. Thus, we have

$$E[y(t)] \approx E[g(z)] + s(t)E[g'(z)] = s(t)E[g'(z)], \quad (3)$$

$$\text{var}[y(t)] = E[y^2(t)] - E[y(t)]^2 \approx E[g^2(z)], \quad (4)$$

where $E[\cdots] = \int_{-\infty}^{\infty} \cdots f_z(z) dz$. Here, for an even PDF f_z and the odd function g , $E[g(z)] = 0$, $E[g(z)g'(z)] = 0$ [6]. The higher order infinitesimal $s^2(t)\{E[g'^2(z)] - E^2[g'(z)]\}$ is neglected [6], resulting in Eq. (4).

The input SNR at $x(t)$ can be defined as the power contained in the spectral line $1/T$ divided by the power contained in the noise background in a small frequency bin ΔB around $1/T$ [7], this is

$$R_{\text{in}} = \frac{|\langle s(t) \exp[-i2\pi t/T] \rangle|^2}{\sigma_z^2 \Delta B \Delta t}, \quad (5)$$

with Δt indicating the time resolution or the sampling period in a discrete-time implementation and the temporal average defined as $\langle \cdots \rangle = \frac{1}{T} \int_0^T \cdots dt$ [7]. Since $s(t)$ is periodic, $y(t)$ is in general be a cyclostationary random signal with period T [7]. Similarly, the output SNR at $y(t)$ is given by

$$R_{\text{out}} = \frac{|\langle E[y(t)] \exp[-i2\pi t/T] \rangle|^2}{\langle \text{var}[y(t)] \rangle \Delta B \Delta t}, \quad (6)$$

with nonstationary expectation $E[y(t)]$ and nonstationary variance $\text{var}[y(t)]$ [7]. Here, we assume the sampling time $\Delta t \ll T$ and observe the output $y(t)$ for a sufficiently large time interval of NT ($N \gg 1$) [7]. Substituting Eqs. (3) and (4) into Eq. (6) and noting Eq. (5), we have

$$R_{\text{out}} \approx \frac{|\langle s(t) \exp[-i2\pi t/T_s] \rangle|^2 E[g'(z)]}{\Delta B \Delta t E[g^2(z)]} = R_{\text{in}} \sigma_z^2 \frac{E[g'(z)]}{E[g^2(z)]}.$$

Thus, the output-input SNR gain G_R of Eq. (1) is

$$G_R = \frac{R_{\text{out}}}{R_{\text{in}}} \approx \sigma_z^2 \frac{E[g'(z)]}{E[g^2(z)]} \leq \sigma_z^2 E \left[\frac{f_z'^2(z)}{f_z^2(z)} \right] = \sigma_z^2 I(f_z), \quad (7)$$

with the equality occurring as g becomes a LOP, viz.

$$g(z) = C f_z'(z)/f_z(z) \triangleq g_{\text{opt}}(z), \quad (8)$$

by the Schwarz inequality for a constant C and $f_z'(z) = df_z(z)/dz$ [2, 5, 6]. It is noted that the LOP g_{opt} of Eq. (8) is odd and accords with the above assumption. More interestingly, the expectation $E[f_z'^2(z)/f_z^2(z)]$ in Eq. (7) is just the FI $I(f_z)$ of the even noise PDF f_z [2, 5]. Furthermore, for an even standardized PDF f_{z_0} with zero mean and unity variance $\sigma_{z_0}^2 = 1$, the scaled noise $z(t) = \sigma_z z_0(t)$ has its PDF $f_z(z) = f_{z_0}(z/\sigma_z)/\sigma_z$. Since the FI satisfies $I(f_z) = I(f_{z_0})/\sigma_z^2$ [1, 8], the output-input SNR gain G_R of Eq. (7) is upper bounded by the FI $I(f_{z_0})$ of a standardized PDF $f_{z_0}(z_0)$, viz.

$$G_R \leq I(f_{z_0}), \quad (9)$$

with equality achieved when g takes the LOP g_{opt} of Eq. (8).

2.2 Performance of a LOD for signal detection

Secondly, we observe a data vector $X = \{x_1, x_2, \dots, x_N\}$ composed of N observation components x_n , which might be the white noise z_n or the mixture of a signal s_n plus white noise z_n [2, 5]. Consider a generalized correlated detector

$$T_{GC}(X) = \sum_{n=1}^N g(x_n) s_n > \gamma, \quad (10)$$

with a memoryless nonlinearity g and the decision threshold γ for the hypotheses $H_1 : x_n = s_n + z_n$, otherwise the hypotheses $H_0 : x_n = z_n$ [5]. Also, we assume that the odd function g has zero mean under the even PDF f_z , i.e. $E[g(x)] = 0$, and there exists a finite bound A such that $0 \leq |s_n| \leq A$. In the asymptotic case of $A \rightarrow 0$ and $N \rightarrow \infty$, the test statistic T_{GC} , according to the central limit theorem, converges to a Gaussian distribution with mean $E[T_{GC}|H_0] = 0$ and variance $\text{var}[T_{GC}|H_0] \approx E[g^2(x)] \sum_{n=1}^N s_n^2$ under the null hypotheses H_0 . Using Eqs. (3) and (4), T_{GC} is asymptotically Gaussian with mean $E[T_{GC}|H_1] \approx E[g'(x)] \sum_{n=1}^N s_n^2$ and variance $\text{var}[T_{GC}|H_1] = \text{var}[T_{GC}|H_0]$ under the hypothesis H_1 [2, 5]. Then, given a false alarm probability P_{FA} , the detection probability of the detector of Eq. (10) is expressed as

$$P_D = Q[Q^{-1}(P_{FA}) - \sqrt{D}], \quad (11)$$

with $Q(x) = \int_x^\infty \exp[-t^2/2]/\sqrt{2\pi} dt$ and its inverse function $Q^{-1}(x)$ [2]. It is seen that P_D is a monotonic increasing function of the deflection coefficient D [2, 5] given by

$$\begin{aligned} D &= \frac{(E[d|H_0] - E[d|H_1])^2}{\text{var}[d|H_0]} \approx \frac{E^2[g'(x)]}{E[g^2(x)]} \sum_{n=1}^N s_n^2 \\ &\leq I(f_z) \sum_{n=1}^N s_n^2 = I(f_{z_0}) \sum_{n=1}^N s_n^2 / \sigma_z^2, \end{aligned} \quad (12)$$

with equality being achieved when $g(x) = g_{\text{opt}}(x)$ of Eq. (8) [2, 5]. This result indicates that the asymptotic optimum detector is the LOD established by the Taylor expansion of the likelihood ratio test statistic $\ln[\prod_{n=1}^N f_z(x_n - s_n) / \prod_{n=1}^N f_z(x_n)] \approx \sum_{n=1}^N g_{\text{opt}}(x_n) s_n$ ($C = -1$) in terms of the generalized Neyman-Pearson lemma [2, 5]. Based on the Bayesian criterion, two hypotheses H_0 and H_1 are endowed with prior probabilities P_0 and $P_1 = 1 - P_0$. Similarly, for the weak signal s_n and the sufficiently large N , the test statistic T_{GC} in Eq. (10) has Gaussian distribution and its performance is evaluated by the error probability [2]

$$P_e = P_0 Q\left(\frac{\ln(\frac{P_0}{P_1})}{\sqrt{D}} + \frac{\sqrt{D}}{2}\right) + P_1 Q\left(\frac{\sqrt{D}}{2} - \frac{\ln(\frac{P_0}{P_1})}{\sqrt{D}}\right), \quad (13)$$

which is also a monotonically decreasing function of D and has a minimum as $D = I(f_{z_0}) \sum_{n=1}^N s_n^2 / \sigma_z^2$ of Eq. (12) for $g(x) = g_{\text{opt}}(x)$ [2, 5]. Interestingly, with $D_G = \sum_{n=1}^N s_n^2 / \sigma_z^2$ (called the signal energy-to-noise ratio of the data vector X [2]) achieved by a matched filter as a benchmark [4, 5], the asymptotic relative efficiency (ARE)

$$\xi_d = D / D_G \leq I(f_{z_0}), \quad (14)$$

provides an asymptotic performance improvement of a detector of Eq. (10) over the linear matched filter [4] when both detectors operate in the same noise environment [2, 4, 5]. The equality of Eq. (14) is achieved as $g(x) = g_{\text{opt}}(x)$ [5].

2.3 Correlation for an aperiodic (random) signal transmission

Thirdly, we transmit a known weak aperiodic signal $s(t)$ through the nonlinearity g of Eq. (1) [11]. Here, the signal $s(t)$ is with the average signal variance $E[s^2(t)] = \sigma_s^2 \ll \sigma_z^2$, the zero mean $E[s(t)] = 0$ and the upper bound A ($0 \leq |s(t)| \leq A$). For example, $s(t)$ can be a sample according to a uniformly distributed random signal equally taking values from a bounded interval. The input cross-correlation coefficient of $s(t)$ and $x(t) = s(t) + z(t)$ is defined as [2, 11]

$$\rho_{s,x} = \frac{E[s(t)x(t)]}{\sqrt{E[s^2(t)]} \sqrt{E[x^2(t)]}} = \frac{\frac{\sigma_s}{\sigma_z}}{\sqrt{\frac{\sigma_s^2}{\sigma_z^2} + 1}} \approx \frac{\sigma_s}{\sigma_z}. \quad (15)$$

Using Eqs. (2)–(4), the output cross-correlation coefficient of $s(t)$ and $y(t)$ is given by

$$\rho_{s,y} = \frac{E[s(t)y(t)]}{\sigma_s \sqrt{\text{var}[y(t)]}} \approx \frac{\sigma_s E[g'(z)]}{\sqrt{E[g^2(z)]}}. \quad (16)$$

Then, the cross-correlation gain (CG) G_ρ is given by

$$G_\rho = \frac{\rho_{s,y}}{\rho_{s,x}} \approx \sigma_z \frac{E[g'(z)]}{\sqrt{E[g^2(z)]}} \leq \sqrt{I(f_{z_0})}, \quad (17)$$

which has its maximal value as $g(z) = g_{\text{opt}}(z)$ of Eq. (8).

2.4 Estimating an unknown parameter of a weak signal

Finally, for the N observation components $x_n = s_n(\theta) + z_n$, we assume the signal $s_n(\theta)$ are with an unknown parameter θ . As the upper bound $A \rightarrow 0$ ($0 \leq |s_n| \leq A$), the Cramér-Rao inequality indicates that the mean squared error of any unbiased estimator of the parameter θ is lower bounded by the reciprocal of the FI [1, 2] given by

$$I(\theta) = \sum_{n=1}^N E \left[\left(\frac{\partial \ln f_z(x_n - s_n)}{\partial \theta} \right)^2 \right]$$

$$\begin{aligned}
&\approx \sum_{n=1}^N \mathbb{E} \left[\left(\frac{df_z(z_n)/dz}{f_z(z_n)} \Big|_{z_n=x_n-s_n} \left(-\frac{\partial s_n}{\partial \theta} \right) \right)^2 \right] \\
&= I(f_z) \sum_{n=1}^N \left(\frac{\partial s_n}{\partial \theta} \right)^2 = \frac{I(f_{z_0})}{\sigma_z^2} \sum_{n=1}^N \left(\frac{\partial s_n}{\partial \theta} \right)^2,
\end{aligned} \tag{18}$$

which indicates that the minimum mean square error of any unbiased estimator is also determined by the FI $I(f_{z_0})$ of distribution with a location shift, as $\sum_{n=1}^N \left(\frac{\partial s_n}{\partial \theta} \right)^2 / \sigma_z^2$ is fixed.

Therefore, just as the FI represents the lower bound of the mean squared error of any unbiased estimator in signal estimation, the physical significance of the FI $I(f_{z_0})$ is that it provides an upper bound of the performance for locally optimum processing for the three considered problems.

3 Extreme value of FI of a standardized PDF

Some interesting questions arise: Which type of noise PDF has a minimal or maximal FI $I(f_{z_0})$, and how large is the extreme value of $I(f_{z_0})$? Does the corresponding LOP in Eq. (8) exist for the noise PDF with extreme $I(f_{z_0})$? These questions will be investigated as follows.

3.1 Minimal Fisher information of a standardized PDF

For a standardized even PDF f_{z_0} , we have

$$I(f_{z_0}) = \mathbb{E} \left[\frac{f_{z_0}'^2(z_0)}{f_{z_0}^2(z_0)} \right] \mathbb{E} [z_0^2] \geq \mathbb{E} \left[\frac{f_{z_0}'(z_0)}{f_{z_0}(z_0)} z_0 \right]^2 = 1, \tag{19}$$

with $\mathbb{E} [z_0^2] = \sigma_{z_0}^2 = 1$ and the equality occurring if $f_{z_0}'(z_0)/f_{z_0}(z_0) = cz_0$ for a constant $c \neq 0$. Then, $f_{z_0}(z_0) = \exp[k + cz_0^2/2]$ [2]. In order to be a PDF, $c < 0$ and $\exp(k)$ is the normalized constant [2]. This is a standardized Gaussian PDF $f_{z_0}(z_0) = \exp(-z_0^2/2)/\sqrt{2\pi}$. Contrarily, any standardized non-Gaussian PDF f_{z_0} has the FI $I(f_{z_0}) > 1$, which indicates that the performance (SNR gain, ARE or CG) is certainly larger than unity via a LOP of Eq. (8) for processing a known weak signal [2, 5, 7].

3.2 Maximal Fisher information of a standardized PDF

A standardized generalized Gaussian noise PDF [2]

$$f_{z_0}(z_0) = c_1(\beta) \exp \left[-c_2(\beta) |z_0|^{\frac{2}{1+\beta}} \right], \tag{20}$$

with $c_1(\beta) = \frac{1}{(1+\beta)} \frac{\Gamma(\frac{1}{2}(\frac{3}{2}(1+\beta)))}{\Gamma(\frac{3}{2}(\frac{1}{2}(1+\beta)))}$ and $c_2(\beta) = \left| \frac{\Gamma(\frac{3}{2}(1+\beta))}{\Gamma(\frac{1}{2}(1+\beta))} \right|^{\frac{1}{1+\beta}}$. The FI $I(f_{z_0})$ of Eq. (20) becomes [2]

$$I(f_{z_0}) = \frac{4}{(1+\beta)^2} \frac{\Gamma[\frac{3}{2}(1+\beta)] \Gamma(\frac{3}{2} - \frac{1}{2}\beta)}{\Gamma^2[\frac{1}{2}(1+\beta)]}, \tag{21}$$

with the corresponding normalized LOP [2]

$$g_{\text{opt}}(x) = |x|^{(1-\beta)/(1+\beta)} \text{sign}(x), \tag{22}$$

where $\text{sign}(\cdot)$ is the sign or signum function. The curve of $I(f_{z_0})$ versus β (cf. Fig. 10.10 of Ref. [2]) clearly indicates that, for $\beta = 0$, $I(f_{z_0}) = 1$ is the minimum corresponding to the standardized Gaussian PDF. It is also noted that, as $\beta \rightarrow 3$ or -1 , $I(f_{z_0}) \rightarrow +\infty$. Is the maximal $I(f_{z_0})$ infinite

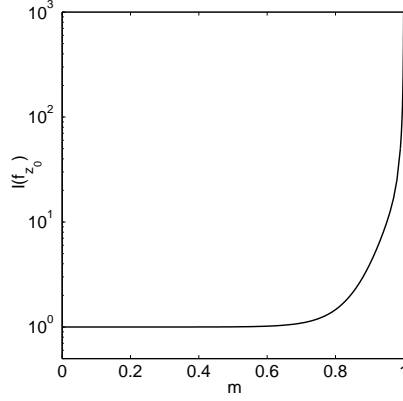


Figure 1: FI $I(f_{z_0})$ versus parameter m of the Gaussian mixture noise PDF of Eq. (26).

for $\beta = 3$ and $\beta = -1$ or not, and is the corresponding LOP simply implemented? The answer is negative, because the LOP of Eq. (22) is not realizable as $g_{\text{opt}}(x) = \pm\infty$ for $|x| > 1$ and $\beta = -1$. When $\beta > 1$, the LOP of Eq. (22) has a singularity at $x = 0$. In this sense, an arbitrary large FI cannot be reached for the generalized Gaussian noise given in Eq. (20).

Next, we consider Gaussian mixture noise $z(t)$ with its PDF

$$f_z(z) = \frac{1}{2\sqrt{2\pi\epsilon^2}} \left[\exp\left(-\frac{(z-\mu)^2}{2\epsilon^2}\right) + \exp\left(-\frac{(z+\mu)^2}{2\epsilon^2}\right) \right], \quad (23)$$

with variance $\sigma_z^2 = \mu^2 + \epsilon^2$ and parameters $\mu, \epsilon \geq 0$. Note that Eq. (23) has another expression [6] as

$$f_z(z) = \exp[-y(z)] / \sqrt{2\pi\epsilon^2}, \quad (24)$$

with $y(z) = \frac{z^2 + \mu^2}{2\epsilon^2} - \ln[\cosh(\frac{\mu z}{\epsilon^2})]$. Based on Eq. (24), the corresponding normalized LOP can be expressed as

$$g_{\text{opt}}(x) = x - \mu \tanh\left(\frac{\mu x}{\epsilon^2}\right). \quad (25)$$

For $0 \leq m \leq 1$, assume $\mu = m\sigma_z$ and $\epsilon^2 = (1-m^2)\sigma_z^2$, Eq. (24) becomes a standardized Gaussian mixture PDF [9]

$$f_{z_0}(z_0) = \exp[-y(z_0)] / \sqrt{2\pi(1-m^2)}, \quad (26)$$

with $y(z_0) = \frac{z_0^2 + m^2}{2(1-m^2)} - \ln[\cosh(\frac{mz_0}{1-m^2})]$. The function of FI $I(f_{z_0})$ versus m is shown in Fig. 1, and $I(f_{z_0})$ can be calculated as (no explicit expression exists)

$$I(f_{z_0}) = \mathbb{E} \left\{ \left[\frac{z_0}{1-m^2} - \frac{m}{1-m^2} \tanh\left(\frac{mz_0}{1-m^2}\right) \right]^2 \right\}. \quad (27)$$

Interestingly, Fig. 1 shows that, as $m = 0$, Eq. (26) is the standardized Gaussian PDF with the FI $I(f_{z_0}) = 1$. While, $I(f_{z_0}) \rightarrow +\infty$ as $m \rightarrow 1$. In Eq. (23), for $m \rightarrow 1$, $\mu = m\sigma_z \rightarrow \sigma_z$ and $\epsilon^2 = (1-m^2)\sigma_z^2 \rightarrow 0$, the term $\lim_{m \rightarrow 1} \frac{1}{\sqrt{2\pi\epsilon^2}} \exp[-\frac{(z-\mu)^2}{2\epsilon^2}] = \delta(z)$ and $\delta(\cdot)$ is the Dirac delta function. Then, Eq. (23) becomes

$$f_z(z) = \frac{1}{2} [\delta(z - \sigma_z) + \delta(z + \sigma_z)], \quad (28)$$

which represents the PDF of dichotomous noise $z(t)$. We also note $\lim_{\epsilon \rightarrow 0, \mu \rightarrow \sigma_z} \tanh\left(\frac{\mu x}{\epsilon^2}\right) = \text{sign}(x)$ ($\mu > 0$) in Eq. (25), and the normalized LOP for dichotomous noise $z(t)$ is

$$g_{\text{opt}}(x) = x - \sigma_z \text{sign}(x). \quad (29)$$

Here, the normalized LOP g_{opt} is not continuous at $x = 0$, but the above analysis is valid for processing a known weak signal in dichotomous noise. This point is like the LOP $g_{\text{opt}}(x) = \text{sign}(x)$ for Laplacian noise [2].

When $z(t)$ randomly takes two levels $-\sigma_z$ and $+\sigma_z$ and $s(t)$ is weak compared with $z(t)$ ($\sigma_z > |s(t)|$), the signs of input data $x(t) = s(t) + z(t)$ always take the sign of $z(t)$, i.e. $\text{sign}(x) = \text{sign}(z)$ in Eq. (29). Therefore, the LOP of Eq. (29) at a fixed time t can be solved as $g[x(t)] = x(t) - \sigma_z \text{sign}[x(t)] = s(t) + z(t) - \sigma_z \text{sign}[z(t)] = s(t)$. Moreover, Refs. [7, 10] have pointed out that there exists a scheme allowing a perfect recovery of $s(t)$ corrupted by dichotomous noise $z(t)$ with the PDF of Eq. (28). However, a practical difficulty in Eq. (22) is that the RMS σ_z needs to be known. The above analysis indicates that the LOP of Eq. (29) can recover the weak signal $s(t)$ perfectly as $\sigma_z > |s(t)|$. Thus, according to the optimum performance of the LOP of Eq. (9) (Eq. (14) or Eq. (17)), the FI $I(f_{z_0}) = \infty$ contained in the type of PDF of Eq. (28), as shown in Fig. 1. Using Eq. (24), the FI $I(f_{z_0})$ of Eq. (27) can be computed as

$$\begin{aligned} I(f_{z_0}) &= \text{E} \left[\left(\frac{dy(z_0)}{dz_0} \right)^2 \right] = \text{E} \left[\frac{d^2 y(z_0)}{dz_0^2} \right] \\ &= \lim_{m \rightarrow 1} \int_{-\infty}^{\infty} \frac{[1 - m^2 - m^2 \text{sech}^2(\frac{mz_0}{1-m^2})]}{(1-m^2)^2} \frac{\exp[-y(z_0)]}{\sqrt{2\pi(1-m^2)}} dz_0 = \infty, \end{aligned} \quad (30)$$

where $\lim_{m \rightarrow 1} [m^2 \text{sech}^2(\frac{mz_0}{1-m^2})] = 0$, the numerator is the infinitesimal $O(1-m^2)$ and the denominator is a higher-order infinitesimal $O((1-m^2)^2)$ in the integral.

4 Conclusions

In this paper, for a known weak signal in additive white noise, it was theoretically demonstrated that the optimum performance of a LOP can be completely determined by the FI of the corresponding standardized even noise PDF, as illustrated by three signal processing case studies: (i) the maximum SNR gain for a periodic signal; (ii) the optimal ARE for signal detection; (iii) the best CG for signal transmission. Thus, our study of the performance of a LOP focused on the measure of the FI of a standardized noise PDF. It is well known that the minimal FI is unity for a standardized Gaussian noise PDF, and the matched filter is the corresponding optimal processor. While, for any non-Gaussian noise, the FI and hence the optimum performance of the LOP is certainly larger than unity. Illustratively, we observed that the generalized Gaussian noise PDF and the Gaussian mixture noise PDF have an arbitrary large FI. There are some types of noise PDF possessing an infinite FI, such as uniform noise and dichotomous noise. However, we argue that only if the LOP is practically realizable, can the performance predicted by the FI be reached in practice. In this sense, it is found that the dichotomous noise has an infinite FI and also a simple LOP structure can be realized in practice.

Some interesting questions arise. For instance, it is known that for a weak signal already corrupted by initial additive white noise, there is usually a LOP that yields the maximal output-input gain. Therefore, can the method of adding an extra amount of noise [9] to the initial data improve the performance of the updated LOP for the resulting noise PDF? This interesting topic will invoke the stochastic resonance phenomenon [6, 7, 9, 11, 12]. Another important question is the influence of finite observation time on the performance of locally optimum processing [9].

References

- [1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley: New York, 1991.
- [2] S. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1998.
- [3] J. Capon, “On the asymptotic efficiency of locally optimum detectors,” *IRE Transactions on Information Theory*, vol. IT-7, pp. 67–71, Apr. 1961.
- [4] N. H. Lu, and B. Eisenstein, “Detection of weak signals in non-Gaussian noise,” *IEEE Transactions on Information Theory*, vol. IT-27(6), pp. 755–771, Nov. 1981.
- [5] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*, New York: Springer-Verlag, 1988.
- [6] S. Zozor and P. O. Amblard, “Stochastic resonance in locally optimal detectors,” *IEEE Transactions on Signal Processing*, vol. 51, pp. 3177–3181, Dec. 2003.
- [7] F. Chapeau-Blondeau, and X. Godivier, “Theory of stochastic resonance in signal transimission by static nonlinear systems,” *Physical Review E*, vol. 55(2), pp. 1478–1495, Feb. 1997.
- [8] A. J. Stam, “Some inequalities satisfied by the quantities of information of Fisher and Shannon,” *Information and Control*, vol. 2, pp. 101–112, 1959.
- [9] F. Chapeau-Blondeau and D. Rousseau, “Raising the noise to improve performance in optimal processing,” *Journal of Statistical Mechanics: Theory and Experiment*, art. no. P01003, 2009.
- [10] M. DeWeese and W. Bialek, “Information flow in sensory neurons,” *Nuovo Cimento*, vol. 17D, pp. 733–741, 1995.
- [11] J. J. Collins, C. C. Chow & T. T. Imhoff, “Stochastic resonance without tuning,” *Nature*, vol. 376, pp. 236–238, 1995.
- [12] S. Kay, “Can detectabilty be improved by adding noise?” *IEEE Signal Processing Letters*, vol. 7, pp. 8–10, Jan. 2000.